The Influence of Autocorrelated Errors on the Bias of Multilevel Time Series Parameter Estimates

I. O. Azeez, R. A. Ipinyomi

Abstract— **The validity of inferences drawn from statistical test results depends on how well data meet associated assumptions. In a two-level multilevel time series model, the standard assumption that the within-individual (level-1) residuals are uncorrelated are rarely checked or little information tends to be reported on whether the data satisfy the assumption underlying the statistical techniques used. Using a simulation approach, the consequences of violating the level-1 independence of observations assumption on the parameter estimates of fixed effects and the associate errors due to bias was investigated. It was found that bias which is generally high, increases with increase autocorrelated errors, and Full maximum likelihood (FML) estimates are more biased than Restricted maximum likelihood (REML) estimates.**

Index Terms— **autocorrelation, multilevel model, repeated measures, simulation.**

I. INTRODUCTION

Many longitudinal studies are designed to investigate changes over time in a characteristic which is measured repeatedly for each study participant, such as clinical trial in which patients are randomly assigned to different treatments and repeatedly evaluated over the course of the study. When measurements are repeated on the same subjects e.g. animals or students, a 2-level hierarchy is established with measurement repetitions or occasions as level 1 units and subjects as level 2 units. In most cases, the multiple observations are taken over time, but they could be over space, such data are referred to as 'repeated measures' or clustered data . A multilevel problem concerns a population with a hierarchical structure. Multilevel models (MLM) were designed to analyze repeated measures data generated from a hierarchical structure and the analysis of such data can be conducted efficiently using a two-level multilevel model.

In some cases especially where measurements are made close together in time, often the error term is not independent through time. Instead, the errors are serially correlated or autocorrelated. If the error term is autocorrelated, the efficiency of ordinary least-squares (OLS) parameter estimates is adversely affected and standard error estimates are biased due to failure to account for the correlated structure of observations. In this paper, we assume data on different subjects are independent, and for simplicity, we assume there are t measurements at the same equally spaced times on each subject.

Multilevel models provide a more accurate and comprehensive description of relationships in clustered data than do conventional models, by correcting underestimated standard errors, by estimating components of variance at several levels, and by estimating cluster-specific intercepts and slopes [7]. The price of such a powerful model for treating hierarchically structured data is the requirement of a set of strong mathematical assumptions whose conditions are expected to be violated to some degree in actual studies. As with other statistical techniques, the assumptions of MLM must be valid in order for the estimates and associated significance tests to have the desired properties.

Multilevel models can accommodate nonindependence of observations, a lack of sphericity, missing data, small and/or discrepant group sample sizes, and heterogeneity of variance across repeated measures [13]. As with most statistical models, an important assumption of MLM is that the level-1 errors (e_{ij}) are independently and normally distributed with a mean of 0 and a variance of σ^2 [12]. This applies to any level-1 model using continuous outcome variables. Mixed linear models are used with repeated measures data to accommodate the fixed effects of covariates and the covariation between observations on the same subject at different times [9]. One of the main reasons we moved to mixed models rather than just working with linear models was to resolve non-independencies in our data, also Linear mixed models provide a powerful and flexible tool for the analysis of a broad variety of data including multilevel data. However, mixed models can still violate independence.

II. INFERENTIAL SETTINGS

In ordinary regression analysis, in the case of severe violations, a variety of statistical methods for correcting nonindependence according to Garson [5] include analysis of variance and other general linear model (GLM) methods that have been adapted to handle non-independence, but these adaptations are problematic. In estimating model parameters when there are random effects, it is necessary to adjust for the covariance structure of the data. The adjustment made by GLM assumes uncorrelated error (that is, it assumes data independence) [5]. Another method for correcting autocorrelation include modeling the serial correlation explicitly using some error autocorrelation formulation, say an Auto Regressive order 1 (AR(1)) process, and then use the generalized least square (GLS) to estimate the Autocorrelation-Corrected [1] .

In multilevel models, specification assumptions apply at each level. Moreover, misspecification at one level can affect results at other levels. In most multilevel applications, the errors in the level-1 model are assumed to have equal variance, σ^2. According to Raudenbush & Bryk [10], if the level-1 variance varies randomly over level-2 units, but these variances are assumed equal, consequences for inference about the level-2 coefficients will be mild, on the other hand if the variances depend systematically as a function of level-1 or

O. I. Azeez, Department of Mathematics and Statistics, Federal Polytechnic, Offa, Kwara State, Nigeria.
R. A. Ipinyomi, Department of Statistics, University of Ilorin, Ilorin, Kwara State, Nigeria.

level-2 predictors, consequences may be more serious. Because causes of heterogeneity are quite different in their implications, it is strongly advocated to investigate possible sources of heterogeneity and model it if found.

Finally, it must be emphasized that failure to adequately account for correlation among repeated measures can result in misleading inferences. For instance, if it is assumed that the repeated measures are uncorrelated when in fact there is strong positive correlation, the nominal standard errors (resulting from the naive assumption of independence or uncorrelated repeated measures) will be incorrect [4]. Autocorrelated data are very common for time ordered data, hence, statistical analysis of repeated measures data must address the issue of covariation between measures on the same unit. A key argument being made is that a systematic study investigating the effects of this violation is important and, therefore, addressed in this paper.

The main question to be answered in this paper is, what is the effect of error due to bias on the efficiency of maximum likelihood (ML) parameters estimates as a result of autocorrelation. Related questions are whether or not the severity of this effect is influenced by the number of measurement occasions, the degree of autocorrelation and the number of subjects. The first two conditions are chosen because when the model includes both random intercepts and slopes (or randomly varying coefficients for any functions of time), the variability of the response can change as a function of the times of measurement, and the magnitudes of the correlations between measurements from the same individual can depend on the time between them.

III. MODEL CONCEPTS

Consider a simple linear regression model for the measurement Y of individual i $(i = 1, 2, \cdots, N \text{ subjects})$ on occasion j $(j = 1, 2, \cdots n \text{ occasions},)$

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + e_{ij} \qquad (1)$$

Ignoring subscripts, this model represents the regression of the outcome variable Y on the independent variable time (denoted X). The subscripts keep track of the particulars of the data, namely whose observation it is (subscript i) and when was this observation made (the subscript j) . The independent variable X gives a value to the level of time, and may represent time in weeks, months, etc. Since Y and X carry both i and j subscripts, both the outcome variable and the time variable are allowed to vary by individuals and occasions.

In linear regression models, like "(1)," the errors e_{ij} are assumed to be normally and independently distributed in the population with zero mean and common variance σ^2. This independence assumption makes the model given in "(1)," an unreasonable one for repeated measure data. This is because the outcomes Y are observed repeatedly from the same individuals, and so it is much more reasonable to assume that errors within an individual are correlated to some degree. Furthermore, the above model posits that the change across time is the same for all individuals since the model parameters (β_0, the intercept or initial level, and β_1, the linear change across time) do not vary by individuals. For both of these reasons, it is useful to add individual-specific effects into the model that will account for the data dependency and describe differential time trends for different individuals. This is precisely what Multilevel Time Series Models for repeated data do.

Estimations of repeated measures data are facilitated by using a multi-level model approach, which allows the estimation of within-individual (level-1) and between-individual (level-2) variations in outcomes. At first, we established a regression equation for the first level variables, in which the tracking results that came from different observation times were the first layer and the invariant individual characteristics were the second layer data.

In the first floor of the data structure, the track observation result was considered as the dependent variable.

$$Y_{ij} = \beta_{oj} + \beta_{1j} X_{ij} + e_{ij} \qquad (2)$$

In a two-level model each term has two subscripts, the first of which corresponds to level 1 while the second refers to level 2. As in (2), subscript "0" means intercept, subscript "1" means slope, subscript "i" means the i-th observation object, Subscript "j" indicates the j- th observation time.

"β_{oj}" is the intercept of the equation, it indicates the average of the i-th observed objects.

"β_{1j}" is the regression coefficient, it indicates the changing rate of the i-th observation object.

"X_{ij}" means the values of the variable X when the i-th observed object is in the j- th observation time.

"e_{ij}" means residual, the implication is that the measured value Y of the i-th object in the j- th observation time that cannot be explained by the independent variable X.

Equation (2) is similar to the general regression equation, the only difference is, intercept and slope are not constant.

In the second layer of the data structures, the intercept and slope are used as the dependent variable in (2), and individual characteristics are considered as independent variables, then we create two regression equations for the second layer:

$$\begin{cases} \beta_{oj} = \gamma_{00} + \gamma_{01} W_j + u_{0j} & (3) \\ \beta_{1j} = \gamma_{10} + \gamma_{11} W_j + u_{1j} & (4) \end{cases}$$

where $$\begin{cases} Y_{ij} = \beta_{0j} + e_{ij} \\ \beta_{oj} = \gamma_{00} + u_{0j} \end{cases} \qquad (5)$$

is referred to as the null model.

In equations (3) and (4), each parameter has two subscripts, if the first subscripts is "0", this is the parameter that relates to the intercept of (2). if the first subscript is "1", this is the parameter that relates to the slope of (2). if the second subscripts is "0", it means the intercept part of the second layer equation, if the second subscript is "1", it means the slope part of the second layer equation.

γ_{00} is the intercept of (3), it can be understood as the average of the dependent variable Y when the independent variable W_j is 0.

W_j is the value on the level-2 predictor

γ_{01} is the regression coefficients of the variables W_j in (3), it can be understood as the impact of the variable W_j to the initial value of the dependent variable Y.

γ_{10} is the intercept of (4), it can be understood as the changing rate of observed object when the variable W_j is 0.

γ_{11} is the regression coefficient of the variable W_j in (4), it can be understood as the effect of the variable W_j on the changing rate.

u_{0j} is the residual of (3), is the intercept deviation for subject i, it represents the influence of individual i on his or her repeated observations.

u_{1j} is the residual of (4) is the slope deviation for subject i

The assumption regarding the independence of the errors is one of conditional independence, that is, they are independent conditional on u_{0j} and u_{1j}.

Our model (2) with one time-level and one individual level explanatory variable can be written as a single complex multilevel time series regression equation by

Substituting $\gamma_{00} + \gamma_{01} W_j + u_{0j}$ for β_{0j}, and Substituting $\gamma_{10} + \gamma_{11} W_j + u_{1j}$ for β_{1j}, and redistributing, we have:

$$Y_{ij} = \gamma_{00} + \gamma_{01} W_j + u_{0j} + \gamma_{10} X_{ij} + \gamma_{11} W_j X_{ij} + u_{1j} X_{ij} + e_{ij} \qquad (6)$$

Rearranging, so that the fixed effects appear first, followed by the random effects, leads us to our final mixed model, defined as

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} W_j + \gamma_{11} X_{ij} W_j + u_{1j} X_{ij} + u_{0j} + e_{ij} \qquad (7)$$

Gill [6] remarks that " In order to allow for the classification of variables and coefficients in terms of the level of hierarchy they affect, a combined model is created by rearranging so that the fixed effects appear first, followed by the random effects.

The term $X_{ij} W_j$ is an interaction term that appears in the model because of modeling the varying regression slope β_{1j} of the time-level variable X_{ij} with the individual level variable W_j.

In equation (7), the errors are no longer independent across the level units. The terms u_{0j} and u_{1j} demonstrate that there is dependency among the level-1 units nested within each level-2 unit. Furthermore, u_{0j} and u_{1j} may have different values within level-2 units, leading to heterogeneous variances of the error terms [12]. That is (7) shows that the composite error structure, $u_{1j} X_{ij} + u_{0j} + e_{ij}$ is now clearly heteroscedastic since it is conditioned on level of the explanatory variable.

IV. METHOD

The simulation model and procedure

We use two different simple two-level models, with one explanatory variable each at the individual level and one explanatory variable at the subject level, conforming to equation (7) above. The model used in the process of generating data for the present study is the first model shown below with W replaced by Z, which henceforth will be referred to as autocorrelated model..

Level-1: $Y_{ij} = \beta_{oj} + \beta_{1j} X_{ij} + e_{ij}$ (2 repeated)

Level-2: $\beta_{0j} = \beta_0 + \beta_{01} Z_j + u_{0j}$ (8)

$\beta_{1j} = \beta_1 + \beta_{11} Z_j + u_{1j}$ (9)

$e_{ij} \sim N(0, \sigma^2 \Omega_k)$.

Thus, $V(y_i) = Z_i \Sigma_v Z_i' + \sigma^2 \Omega_k$ (10)

where Ω_k depends on q autocorrelation parameters, with q varying depending on the type of autocorrelated error structure being considered.

The motivation lies in the need to allow for patterns of dependence, rather than complete independence among response values. The simplest way to allow such dependence is to assume $V(y_i) = Z_i \Sigma_v Z_i' + \sigma^2 \Omega_k$ with Ω_k of dimension $N \times N$, symmetric and positive definite or semi positive definite (which allows any covariance matrix).

The second model is given below, which will henceforth be referred to as standard model.

Level-1: $Y_{ij} = \beta_{oj} + \beta_{1j} X_{ij} + e_{ij}$ (2 repeated)

Level-2: $\beta_{0j} = \beta_0 + \beta_{01} Z_j + u_{0j}$ (8 repeated)

$\beta_{1j} = \beta_1 + \beta_{11} Z_j + u_{1j}$ (9 repeated)

$e_i \sim N(0, \sigma^2 I)$.

Thus, $V(y_i) = Z_i \Sigma_v Z_i' + \sigma^2 I$ (11)

V. ESTIMATION METHODS FOR VARIANCE COMPONENTS

For the purpose of this study, used were made of R program to estimate the parameters. The estimation methods are compared in relation to the number of subjects, number of measurement occasion, and autocorrelation coefficient under the following conditions:

I. autocorrelation coefficients of 0.3, 0.7, 0.99
II. variances of intercept and slopes and their covariances of − 12.63, 2.08 and -1.42 respectively
III. numbers of subjects − 30, 50, and 100
IV. numbers of observation within subjects − 3, 5, and 10
V. 1000 replication for each condition
   - For the regression coefficients, 1.00 was chosen for the intercept, and 0.3 for all the regression slopes [2] [8]. The first level variance σ^2 was fixed at 12.22, while the error terms in the simulated data are auto regressively correlated. The sizes of the conditions are partially based on literature and partially on practical experience.

VI. RESULTS

Coverage

In order to investigate the influence of the number of subjects, the autocorrelation coefficients and the number of measurement occasions on the estimation of error of bias on the parameter estimates, the coverage per condition was calculated to describes the uncertainty inherent in our estimate, and describes a range of values within which we can be reasonably sure that the true effect actually lies.

Wald simplest 95% confidence intervals (CI) on the estimated average slopes were constructed, by taking the point estimate ±1.96 estimated standard errors in order to determine lopsidedness of coverage resulting from errors due to bias and find the influence of the number of subjects, the autocorrelation coefficient and the number of measurement occasions on the constructed CI for the parameter estimates. More specifically, when using the Wald Confidence Interval, two points on either side of MLE are chosen such that they are equidistant from MLE value (MLE ± SE * (1-alpha)/2 percentile of Normal distribution).

The width of REML confidence interval estimates are wider than the ML confidence interval, but the differences are small. In standard multi-level model, combinations with same

numbers of subjects and measurements occasion appear to have the same width as their prediction interval, this adds quite a bit to our understanding of the variability in our random coefficients. Also observed from the constructed confidence interval, there is lopsidedness of coverage resulting in estimates falling more frequently to one side than the other of the true parameter.

To substantiate our claim, standard normal distribution was used to estimate the expected percentage of regression coefficients that are less than 0.3 under autocorrelated model (MLΩ), and found to be within the range of 0.15% to 100%. Lopsidedness of coverage is a direct consequence of the bias in the multilevel point estimator, on which the Wald interval is centered. Despite this problem, multilevel Wald 95% intervals appear to provide conservatively valid (i.e. at least 95%) average coverage for the parameter estimates [11].

In our constructed 95% confidence interval, subjects (level-2 units) with low numbers of measurement occasion and low autocorrelation coefficients are predicted to have a wider confidence interval than subjects with high numbers of measurement occasion and high autocorrelation coefficient. Similarly, differences between the number of subjects indicate relationship between the width of the confidence interval. Not surprising, intercept and slope coefficients are random variables that vary across subjects, the specific values for the intercept and slope coefficients are subjects characteristics.

VII. EVALUATING BIAS

For the assessment of the parameter estimates, the absolute bias was considered for each parameter.

Table I. Comparing the bias of slopes for measurement occasion, t = 3 for two estimators under standard and autocorrelated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope.

	ML (Ω)			**REML (Ω)**			**ML ($\boldsymbol{I}$)**			**REML ($\boldsymbol{I}$)**		
N	ρ			ρ			ρ			ρ		
	0.3	0.7	0.99	0.3	0.7	0.99	0.3	0.7	0.99	0.3	0.7	0.99
30	-0.6267	-0.7429	-0.7843	-0.6172	-0.7328	-0.7740	-0.4629	-0.4106	-0.3726	-0.4257	-0.3734	-0.3354
50	0.3721	0.4295	0.4654	0.3838	0.4407	0.4765	0.3542	0.3775	0.3943	0.3578	0.3811	0.3980
100	0.5482	0.6212	0.6683	0.5575	0.6304	0.6775	0.5300	0.5781	0.6130	0.5229	0.5710	0.6060

Where $\boldsymbol{\Omega}$ represents autocorrelation matrix and $\boldsymbol{I}$ is the identity matrix.

Table II. Comparing the bias of slopes for measurement occasion, t = 5 for two estimators under standard and autocorrelated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope.

	ML (Ω)			**REML (Ω)**			**ML ($\boldsymbol{I}$)**			**REML ($\boldsymbol{I}$)**		
N	ρ			ρ			ρ			ρ		
	0.3	0.7	0.99	0.3	0.7	0.99	0.3	0.7	0.99	0.3	0.7	0.99
30	-0.5107	-0.6591	-0.6925	-0.5183	-0.6653	-0.6987	-0.3521	-0.3670	-0.3778	-0.3414	-0.3563	-0.3671
50	0.1776	0.3169	0.3377	0.1649	0.3050	0.3260	-0.0257	-0.0290	-0.0314	-0.0293	-0.0326	-0.0350
100	0.8363	0.9759	0.9979	0.8347	0.9734	0.9952	0.6397	0.6402	0.6405	0.6326	0.6331	0.6335

Where $\boldsymbol{\Omega}$ represents autocorrelation matrix and $\boldsymbol{I}$ is the identity matrix.

Table III. Comparing the bias of slopes for measurement occasion, t = 10 for two estimators under standard and autocorrelated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope.

	ML (Ω)			**REML (Ω)**			**ML ($\boldsymbol{I}$)**			**REML ($\boldsymbol{I}$)**		
N	ρ			ρ			ρ			ρ		
	0.3	0.7	0.99	0.3	0.7	0.99	0.3	0.7	0.99	0.3	0.7	0.99
30	-0.2669	0.0363	0.1196	-0.2699	0.0314	0.1141	-0.4884	-0.4770	-0.4687	-0.4930	-0.4816	-0.4733
50	0.9309	1.3538	1.4423	0.9259	1.3475	1.4356	0.5867	0.5901	0.5925	0.5855	0.5889	0.5913
100	0.2546	0.3573	0.3771	0.2519	0.3531	0.37254	0.1663	0.1656	0.1651	0.1671	0.1664	0.1659

Where $\boldsymbol{\Omega}$ represents autocorrelation matrix and $\boldsymbol{I}$ is the identity matrix.

.

VIII. COMMENT:

As can be seen in Table I to Table III, the random estimator with or without autocorrelated errors, generally show a large bias for small numbers of repeated measures and generally, an increase bias when $\boldsymbol{\rho}$, the autocorrelation coefficient increases. The difference between MLE and RMLE estimates are very small and inconsistent over conditions. Generally, the bias of MLE is larger than the bias of RMLE. Similarly, autocorrelated models exhibit higher bias compared to standard multilevel model. Varying the number of observations within a fixed subject size does not provide a clear indication of neither an increase nor decrease in bias.

As for the amount of bias in the ML parameter estimates of the standard model, we see that the naive two-stage standard model consistently underestimates the true bias when incorrectly assuming compound symmetry.

The observations above are in consistent with theory, Demidenko [3] state that:

- For some (co)variance parameters, when few subjects are sampled, no matter how many observations are sampled per unit, problematic bias remains.
- Complex models exhibit greater bias than simpler models.
- Full maximum likelihood (FML) estimates are more biased than Restricted maximum likelihood (REML) estimates.
- Raudenbush and Bryke [10] reported that RML estimates variance components after removing the fixed effects from the model, can lead theoretically to less bias than FML, especially when the number of groups is small.

IX. DISCUSSION AND CONCLUSION

Bias which is generally high, increases with increase autocorrelated errors The results illustrate the generality of the theorem and the substantial bias that can occur. Even with a correctly specified covariance model, observed bias for smaller sample sizes is large, though consistent with the theory, an indication that the result of shrinkage are most noticeable if the number of observations in single individual is small. It seems the bias *has* a net direction and magnitude so that averaging it over a large number of observations does not eliminate its effect, and Increasing the sample size is not going to help.

In the 95% CI for the slopes, we see that FML slopes have greater *precision* than REML. Another advantage of multilevel models is that they incorporate the precision of estimates into the model. When autocorrelation in the error is considered, heterogeneity was found to be lost (incorrectly) contrary to our usual expectation for random-effects models, where precision will decrease with increasing heterogeneity and confidence intervals will be widen correspondingly. Based on these results, it can be concluded that having autocorrelated errors in the repeated measures data increase the biasness of ML estimates with autocorrelated model exhibiting large bias compared to standard model. Analysing the data ignoring the existing autocorrelated errors mask the effect of error due to bias on the parameter estimates and our estimators (ML) under autocorrelated model, though with large bias, is expected to reduce some loss function (particularly mean squared error) compared with unbiased estimators (since our estimators are shrinkage estimators). Similarly, the FML estimates with large bias are not necessarily less accurate than REML estimates as will be judged by the expected mean square error.

The shorter FML-based intervals resulting from the assumption that the fixed effects in the model are equal to their ML estimates is expected to converge as the number of measurement occasions j becomes large.

REFERENCES

[1] Anya, M. M & Aris, S. (2002). The Linear regression model with autocorrelated errors: Just Say No to Error Autocorrelation. Annual Meeting of the American Agricultural Economics Association, July 28-31, 2002.

[2] Cohen, J. (1988). Statistical power analysis for the behavioral sciences. New York: LEA.

[3] Demidenko, E. (2004). Mixed models: theory and applications. New York: John Wiley & Sons.

[4] Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). Applied longitudinal analysis, Second Edition, Hoboken, New Jersey: John Wiley & Sons, Inc.

[5] Garson David (2013). Hierarchical linear modeling: guide and applications
SAGE publication inc.

[6] Gill, J. (2002). Bayesian methods for the social and behavioral sciences. Chapman & Hall, New York.

[7] Kreft, I. G. G., & De Leeuw, J. (1998). Introducing multilevel modeling. Thousand Oaks, CA: Sage.

[8] Maas, C.J.M., & Hox, J.J. (2005). Sufficient sample sizes for multilevel modeling methodology.

[9] Ramon, C. L., Jane, P. Ranjini, N. (2000). Tutorial in biostatistics modeling covariance structure in the analysis of repeated measures data. statistics in medicine, Statist Med 19:1793{1819}

[10] Raudenbush, S. W. & Bryk, A. S. (2002). Hierarchical linear models: applications and data analysis methods (2nd ed.). Thousand Oaks, CA: Sage Publications.

[11] Sander,G.(2000). Principle of multilevel modeling. International Journal of Epidemiology. Int J Epidemiol (2000) 29 (1): 158-167.

[12] Sullivan, L. M., Dukes, K. A., & Losina, E. (1999). Tutorial in biostatistics: An introduction to hierarchical linear modeling. Statistics in Medicine, 18, 855-888.

[13] Woltmam, H., Andrea, F., Christine, J. M. &, Meredith R. (2012). An introduction to hierarchical linear modeling. Tutorial in quantitative methods for psychology 2012, Vol. 8(1), p. 52-69.

O. I. Azeez, Department of Mathematics and Statistics, Federal Polytechnic, Offa, Kwara State, Nigeria.

R. A. Ipinyomi, Department of Statistics, University of Ilorin, Ilorin, Kwara State, Nigeria.